Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational and Applied Mathematics 206 (2007) 322–340
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

The Hertz contact problem, coupled Volterra integral equations and a linear complementarity problem

A. Gauthier[1], P.A. Knight*, S. McKee

Department of Mathematics, University of Strathclyde, 26 Richmond Street, Glasgow G1 1XH, UK

Received 9 May 2006

Abstract

This paper is concerned with the indentation of an elastic half-space by an axisymmetric punch under a monotonically applied normal force and under the assumption of Coulomb friction with coefficient μ in the region of contact. Within an inner (unknown) circle the contact is adhesive, while in the surrounding annulus the surface moves inwards with increasing load. In this paper it is shown how this problem is equivalent to two coupled Abel's equations with an unknown free point, the inner circumference of the annulus. It is further shown that a product integration finite difference approximation of those integral equations leads to a mixed linear complementarity problem (mixed LCP). A method based on Newton's method for solving non-smooth nonlinear equations is demonstrated to converge under restrictive assumptions on the physical parameters defining the system; and numerical experimentation verifies that it has much wider applicability. The method is also validated against the approach of Spence. The advantage of the mixed LCP formulation is that it provides the radius of the inner adhesive circle directly using the physical parameters of the problem.

MSC: 47B35; 74B10; 90C33

Keywords: Hertz contact problem; Abel's integral equations; Linear complementarity problem

1. Introduction

The elastic contact problem with the Coulomb friction law in the region of contact can be formulated in various ways requiring sometimes very different numerical methods for their solution.

A variational inequality formulation involves stress and displacement with a non-differentiable term arising from the Coulomb friction law for elastostatic problems. It may also be regarded as a minimisation problem with a complementarity constraint (see e.g., [7]). Using a finite element method approximation these formulations have been considered by Kikuchi and Oden [12]. Dynamic contact problems have also been solved using algorithms developed for solving linear complementarity problems (LCPs) (see e.g., [11,1]).

An alternative approach involves stress and displacement on the domain boundary of the contact problem (see e.g., [14]); in this case the biharmonic equation (for the Airy stress function) may be characterised as coupled Abel's

* Corresponding author.
E-mail address: p.a.knight@strath.ac.uk (P.A. Knight).
[1] Current address: Index Education, BP201, 13322, Marseille, Cedex 16, France.

0377-0427/$ - see front matter
doi:10.1016/j.cam.2006.07.008

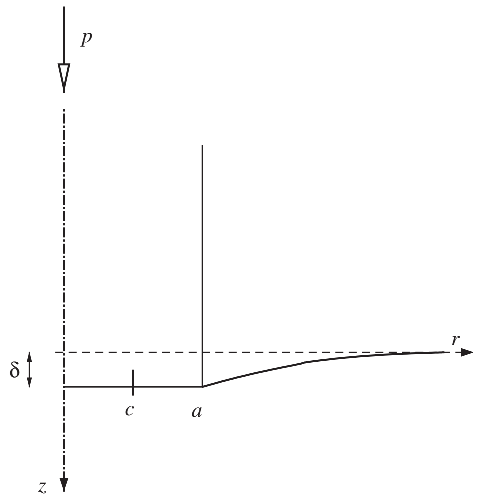

Fig. 1. Flat punch.

equations. This is particularly convenient for certain geometries. The Coulomb friction law can in this case (as in others) be formulated by a complementarity condition involving both stress and displacement.

The paper will be concerned with this integral equation formulation of the punch indentation problem (with partial adhesion) as illustrated in Fig. 1. A monotonic normal force p is applied to a rigid axisymmetric punch of radius a resting on an elastic half-space causing an indentation δ in the vertical direction. The Coulomb friction law with coefficient μ in the region of contact is assumed to be satisfied with no slip occurring over the central circle whose radius is $r = ca$, $0 < c < 1$.

The integral equation formulation was first proposed by Noble and Spence [13]. Later Spence [15] suggested a numerical method for solving the coupled singular Volterra equations with the unknown free point ca (the point separating slip from no slip). Essentially this involved guessing the value of c and iterating until the correct (known) friction coefficient was determined. This might be called an indirect approach.

The object of this paper is to solve the problem directly by solving both for the stress and displacement and for the unknown free point. In Section 2 the system of coupled Abel's equations is derived. In Section 3, after a change of variable, the system is discretised over a regular mesh. The resulting discretisations are comparable to methods proposed in [16,2] for uncoupled systems. Then in Section 4 it is shown that the discrete form is in fact a mixed LCP. An efficient numerical method developed by Chen and Mangasarian [6] is then implemented in Section 5 and in Section 6 convergence of this iterative method is discussed making use of the Toeplitz structure of the associated matrices. Numerical results are presented in Section 7.

The motivation for this work arose from a Study Group with Industry: the UK Ministry of Agriculture, Fisheries and Food (as it was then) had developed a prototype device for measuring the freshness of fish by indenting the surface with a punch and making use of the fact that freshness was directly related to the "degree of elasticity". In reality a fish is not elastic. It might be regarded as viscoelastic or as a fluid-filled porous matrix, and it is recovery time that is the crucial determinant. Thus this work, while providing an efficient numerical method for solving the problem of a punch indenting an elastic surface, does not address the much more difficult problem of the rheology of fish and must therefore, be regarded as preliminary.

2. Derivation of the coupled Abel's equations

This section follows the approach of Noble and Spence [13]. Recall the definition of the Hankel transform

$$\overline{f}_\nu(\xi) = \mathscr{H}[f(x), \xi] = \int_0^\infty x f(x) J_\nu(x\xi)\,\mathrm{d}x.$$

The Hankel inversion formula tells us that if $f(\rho)$ is continuous at the point $\rho = r$ then

$$\int_0^\infty u\overline{f}_\nu(u)J_\nu(ur)\,\mathrm{d}u = f(r),$$

that is, $\mathscr{H}[\overline{f}_\nu(u), r] = f(r)$.

It is well known that the radial and normal displacements and stresses are recoverable from the Airy stress function which itself satisfies a biharmonic equation. By applying the theory of Hankel transforms this equation may be reduced to a fourth order ordinary differential equation which may be readily solved to provide (see [14])

$$u(r,z) = \frac{\lambda+\mu}{\mu}\int_0^\infty t^2\frac{\partial G}{\partial z}J_1(rt)\,\mathrm{d}t,$$

$$w(r,z) = \int_0^\infty t\left(\frac{\partial^2 G}{\partial z^2} - \frac{\lambda+2\mu}{\mu}t^2 G\right)J_0(rt)\,\mathrm{d}t,$$

$$\sigma(r,z) = \int_0^\infty t\left((\lambda+2\mu)\frac{\partial^3 G}{\partial z^3} - (3\lambda+4\mu)t^2\frac{\partial G}{\partial z}\right)J_0(rt)\,\mathrm{d}t,$$

$$\tau(r,z) = \int_0^\infty t^2\left(\lambda\frac{\partial^2 G}{\partial z^2} + (\lambda+2\mu)t^2 G\right)J_1(rt)\,\mathrm{d}t,$$

for the radial and axial displacements, radial and shear stress components, where $\lambda = E\nu/(1-\nu-2\nu^2)$ and $\mu = E/(2+2\nu)$ are the usual Lamé constants (E is Young's modulus and ν is Poisson's ratio); $G(t,z) = [c(t) + zD(t)]\exp(-tz)$, where $c(t)$ and $D(t)$ are unknown functions to be found from boundary conditions; and it has been assumed that the stresses vanish as $z \to \infty$.

To obtain a form convenient for discussing stresses and displacements at a surface, we put $z = 0$ in these equations and make the substitutions $\sigma(r,0) = \sigma(r)$, $\tau(r,0) = \tau(r)$, etc., to give

$$\sigma(r) = -2\mu\int_0^\infty t\phi(t)J_0(rt)\,\mathrm{d}t, \tag{2.1}$$

$$\tau(r) = -2\mu\int_0^\infty t\psi(t)J_1(rt)\,\mathrm{d}t, \tag{2.2}$$

$$w(r) = 2(1-\nu)\int_0^\infty \{\phi(t) - \gamma\psi(t)\}J_0(rt)\,\mathrm{d}t, \tag{2.3}$$

$$u(r) = 2(1-\nu)\int_0^\infty \{\psi(t) - \gamma\phi(t)\}J_1(rt)\,\mathrm{d}t, \tag{2.4}$$

where $\gamma = (1-2\nu)/(2-2\nu)$.

2.1. Relations between stresses and displacements

We now compute displacements in terms of stresses and vice versa. We first note, from (2.1) and (2.2), that $-\sigma(r)/2\mu$ and $-\tau(r)/2\mu$ are the Hankel transforms of the (unknown) functions $\phi(t)$ and $\psi(t)$, that is,

$$\phi(t) = \int_0^\infty r(-\sigma(r)/2\mu)J_0(rt)\,\mathrm{d}r, \tag{2.5}$$

$$\psi(t) = \int_0^\infty r(-\tau(r)/2\mu)J_1(rt)\,\mathrm{d}r. \tag{2.6}$$

It is convenient to introduce modified forms of (2.1)–(2.4).

Let $\alpha(t)=\phi(t)-\gamma\psi(t)$, $\beta(t)=\psi(t)-\gamma\phi(t)$, so that

$$\phi(t)=\frac{\alpha(t)+\gamma\beta(t)}{1-\gamma^2},\quad \psi(t)=\frac{\gamma\alpha(t)+\beta(t)}{1-\gamma^2}$$

and (2.1)–(2.4) become

$$\sigma(r)=-\frac{2\mu}{1-\gamma^2}\int_0^\infty t(\alpha(t)+\gamma\beta(t))J_0(rt)\,\mathrm{d}t,$$

$$\tau(r)=-\frac{2\mu}{1-\gamma^2}\int_0^\infty t(\gamma\alpha(t)+\beta(t))J_1(rt)\,\mathrm{d}t,$$

$$w(r)=2(1-v)\int_0^\infty \alpha(t)J_0(rt)\,\mathrm{d}t,$$

$$u(r)=2(1-v)\int_0^\infty \beta(t)J_1(rt)\,\mathrm{d}t.$$

Using the relations

$$\frac{\mathrm{d}}{\mathrm{d}x}J_0(x)=-J_1(x)\quad\text{and}\quad \frac{\mathrm{d}}{\mathrm{d}x}(xJ_1(x))=xJ_0(x),$$

we obtain

$$\sigma(r)=-\frac{2\mu}{1-\gamma^2}\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\left[r\int_0^\infty(\alpha(t)+\gamma\beta(t))J_1(rt)\,\mathrm{d}t\right], \tag{2.7}$$

$$\tau(r)=\frac{2\mu}{1-\gamma^2}\frac{\mathrm{d}}{\mathrm{d}r}\left[\int_0^\infty(\gamma\alpha(t)+\beta(t))J_0(rt)\,\mathrm{d}t\right], \tag{2.8}$$

$$\frac{\mathrm{d}w(r)}{\mathrm{d}r}=-2(1-v)\int_0^\infty t\alpha(t)J_1(rt)\,\mathrm{d}t, \tag{2.9}$$

$$\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}(ru(r))=2(1-v)\int_0^\infty t\beta(t)J_0(rt)\,\mathrm{d}t. \tag{2.10}$$

We illustrate how this is achieved for (2.7)—the others follow similarly. Let $x=rt$ so that $\mathrm{d}x=t\,\mathrm{d}r$. Thus

$$rtJ_0(rt)=\frac{\mathrm{d}}{\mathrm{d}(rt)}(rtJ_1(rt))=\frac{1}{t}\frac{\mathrm{d}}{\mathrm{d}r}(rtJ_1(rt))=\frac{\mathrm{d}}{\mathrm{d}r}(rJ_1(rt)).$$

That is,

$$tJ_0(rt)=\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}(rJ_1(rt))$$

and the result follows assuming differentiation can be taken outside the integral.

Now (2.9) and (2.10) are Hankel transforms and can be inverted, that is

$$\alpha(t)=-\frac{1}{2(1-v)}\int_0^\infty s\frac{\mathrm{d}w(s)}{\mathrm{d}s}J_1(st)\,\mathrm{d}s, \tag{2.11}$$

$$\beta(t)=\frac{1}{2(1-v)}\int_0^\infty \frac{\mathrm{d}}{\mathrm{d}s}(su(s))J_0(st)\,\mathrm{d}s. \tag{2.12}$$

2.2. Relations between Abel integrals involving the stresses and displacements

Consider (2.5) and take the Fourier cosine transform

$$\begin{aligned}\phi_c(x) &= \int_0^\infty \phi(t)\cos xt\,\mathrm{d}t = -\frac{1}{2\mu}\int_0^\infty\left(\int_0^\infty r\sigma(r)J_0(rt)\,\mathrm{d}r\right)\cos xt\,\mathrm{d}t \\ &= -\frac{1}{2\mu}\int_0^\infty r\sigma(r)\left(\int_0^\infty J_0(rt)\cos xt\,\mathrm{d}t\right)\mathrm{d}r.\end{aligned}$$

However,

$$\int_0^\infty J_0(rt)\cos xt\,\mathrm{d}t = \begin{cases}(r^2-x^2)^{-1/2}, & 0<x<r,\\ 0, & x>r,\end{cases}$$

and so $\phi_c(x)$ becomes

$$\phi_c(x) = -\frac{1}{2\mu}\int_x^\infty \frac{s\sigma(s)}{(s^2-x^2)^{1/2}}\,\mathrm{d}s. \tag{2.13}$$

Similarly

$$\phi_s(x) = \int_0^\infty \phi(t)\sin xt\,\mathrm{d}t = -\frac{1}{2\mu}\int_0^x \frac{s\sigma(s)}{(x^2-s^2)^{1/2}}\,\mathrm{d}s \tag{2.14}$$

using

$$\int_0^\infty J_0(rt)\sin xt\,\mathrm{d}t = \begin{cases}0, & 0<x<r,\\ (x^2-r^2)^{-1/2}, & x>r.\end{cases}$$

Now

$$\begin{aligned}\psi_c(x) &= -\frac{1}{2\mu}\int_0^\infty\left(\int_0^\infty r\tau(r)J_1(rt)\,\mathrm{d}r\right)\cos xt\,\mathrm{d}t \\ &= -\frac{1}{2\mu}\int_0^\infty r\tau(r)\left(\int_0^\infty J_1(rt)\cos xt\,\mathrm{d}t\right)\mathrm{d}r \\ &= -\frac{1}{2\mu}\left[\int_0^x \tau(r)[1-x/(x^2-r^2)^{1/2}]\,\mathrm{d}r + \int_x^\infty \tau(r)\,\mathrm{d}r\right] \\ &= -\frac{1}{2\mu}\left[\int_0^x \tau(r)\,\mathrm{d}r - x\int_0^x \frac{\tau(r)}{(x^2-r^2)^{1/2}}\,\mathrm{d}r + \int_x^\infty \tau(r)\,\mathrm{d}r\right] \\ &= -\frac{1}{2\mu}\left[\int_0^\infty \tau(s)\,\mathrm{d}s - x\int_0^x \frac{\tau(s)}{(x^2-s^2)^{1/2}}\,\mathrm{d}s\right]\end{aligned} \tag{2.15}$$

using

$$\int_0^\infty J_1(rt)\cos xt\,\mathrm{d}t = \begin{cases}r^{-1}, & 0<x<r,\\ r^{-1}[1-x(x^2-r^2)^{-1/2}], & x>r.\end{cases}$$

Similarly, using the fact that

$$\int_0^\infty J_1(rt)\sin xt\,\mathrm{d}t = \begin{cases}x/r(r^2-x^2)^{-1/2}, & 0<x<r,\\ 0, & x>r,\end{cases}$$

$$\psi_s(x) = \int_0^\infty \psi(t)\sin xt\,\mathrm{d}t = -\frac{x}{2\mu}\int_x^\infty \frac{\tau(s)}{(s^2-x^2)^{1/2}}\,\mathrm{d}s. \tag{2.16}$$

Consider now the Fourier cosine and sine transforms of $\alpha(t)$ and $\beta(t)$. From (2.11) and (2.12)

$$
\begin{aligned}
\alpha_c(x) &= -\frac{1}{2(1-v)}\int_0^\infty\left(\int_0^\infty s w'(s) J_1(st)\,\mathrm{d}s\right)\cos xt\,\mathrm{d}t \\
&= -\frac{1}{2(1-v)}\int_0^\infty s w'(s)\left(\int_0^\infty J_1(st)\cos xt\,\mathrm{d}t\right)\mathrm{d}s \\
&= -\frac{1}{2(1-v)}\left(\int_0^x s w'(s) s^{-1}[1-x(x^2-s^2)^{-1/2}]\,\mathrm{d}s + \int_x^\infty s w'(s) s^{-1}\,\mathrm{d}s\right) \\
&= -\frac{1}{2(1-v)}\left(w(x)-w(0)-\int_0^x \frac{x}{(x^2-s^2)^{1/2}} w'(s) + w(\infty) - w(x)\right) \\
&= \frac{1}{2(1-v)}\left(w(0) + x\int_0^x \frac{w'(s)}{(x^2-s^2)^{1/2}}\,\mathrm{d}s\right),
\end{aligned} \tag{2.17}
$$

where here and henceforth $w'(s)$ will denote $(\mathrm{d}w/\mathrm{d}s)(s)$.

In a similar fashion, using the expressions for the Fourier transforms of the Bessel functions $J_0(xt)$ and $J_1(xt)$, we obtain

$$\alpha_s(x) = -\frac{x}{2(1-v)}\int_x^\infty \frac{w'(s)}{(s^2-x^2)^{1/2}}\,\mathrm{d}s, \tag{2.18}$$

$$\beta_c(x) = \frac{1}{2(1-v)}\int_x^\infty \frac{[su(s)]'}{(s^2-x^2)^{1/2}}\,\mathrm{d}s, \tag{2.19}$$

$$\beta_s(x) = \frac{1}{2(1-v)}\int_0^x \frac{[su(s)]'}{(x^2-s^2)^{1/2}}\,\mathrm{d}s. \tag{2.20}$$

Now since $\alpha(t) = \phi(t) - \gamma\psi(t)$ and $\beta(t) = \psi(t) - \gamma\phi(t)$ we may take Fourier transforms of both sides to obtain

$$\phi_c(x) - \gamma\psi_c(x) = \alpha_c(x), \tag{2.21}$$

$$\psi_s(x) - \gamma\phi_s(x) = \beta_s(x), \tag{2.22}$$

$$\psi_c(x) - \gamma\phi_c(x) = \beta_c(x), \tag{2.23}$$

$$\psi_s(x) - \gamma\psi_s(x) = \alpha_s(x). \tag{2.24}$$

Inserting (2.13)–(2.16) into (2.21)–(2.24) results in

$$
\begin{aligned}
&\int_x^\infty \frac{s\sigma(s)}{(s^2-x^2)^{1/2}}\,\mathrm{d}s - \gamma\left(\int_0^\infty \tau(s)\,\mathrm{d}s - x\int_0^x \frac{\tau(s)}{(x^2-s^2)^{1/2}}\,\mathrm{d}s\right) \\
&\quad = \frac{\mu}{1-v}\left(w(0) + x\int_0^x \frac{w(s)}{(x^2-s^2)^{1/2}}\,\mathrm{d}s\right) \\
&\quad = -\frac{\mu}{1-v}\frac{\mathrm{d}}{\mathrm{d}x}\left[\int_0^x \frac{s w(s)}{(x^2-s^2)^{1/2}}\,\mathrm{d}s\right],
\end{aligned} \tag{2.25}
$$

$$
\begin{aligned}
&x\int_x^\infty \frac{\tau(s)}{(s^2-x^2)^{1/2}}\,\mathrm{d}s - \gamma\int_0^x \frac{s\sigma(s)}{(x^2-s^2)^{1/2}}\,\mathrm{d}s \\
&\quad = \frac{\mu}{1-v}\int_0^x \frac{[su(s)]'}{(x^2-s^2)^{1/2}}\,\mathrm{d}s = -\frac{\mu}{1-v}\frac{\mathrm{d}}{\mathrm{d}x}\left[x\int_0^x \frac{u(s)}{(x^2-s^2)^{1/2}}\,\mathrm{d}s\right],
\end{aligned} \tag{2.26}
$$

$$\int_0^\infty \tau(s)\,\mathrm{d}s - x\int_0^x \frac{\tau(s)}{(x^2-s^2)^{1/2}}\,\mathrm{d}s - \gamma\int_x^\infty \frac{s\sigma(s)}{(s^2-x^2)^{1/2}}\,\mathrm{d}s$$
$$= \frac{\mu}{1-v}\int_x^\infty \frac{[su(s)]'}{(s^2-x^2)^{1/2}}\,\mathrm{d}s = -\frac{\mu}{1-v}\frac{\mathrm{d}}{\mathrm{d}x}\left[x\int_x^\infty \frac{u(s)}{(s^2-x^2)^{1/2}}\,\mathrm{d}s\right] \tag{2.27}$$

and

$$\int_0^x \frac{s\sigma(s)}{(x^2-s^2)^{1/2}}\,\mathrm{d}s - \gamma x\int_x^\infty \frac{\tau(s)}{(s^2-x^2)^{1/2}}\,\mathrm{d}s$$
$$= -\frac{\mu x}{1-v}\int_x^\infty \frac{w'(s)}{(s^2-x^2)^{1/2}}\,\mathrm{d}s = \frac{\mu}{1-v}\frac{\mathrm{d}}{\mathrm{d}x}\left[\int_x^\infty \frac{sw(s)}{(s^2-x^2)^{1/2}}\,\mathrm{d}s\right]. \tag{2.28}$$

Eqs. (2.25) and (2.26) are involved in contact problems where the stresses on $r > a$ and the displacements on $r < a$ are given:

$$\begin{cases} \sigma(r) = \tau(r) = 0, & r > a, \\ w(r) = \delta, & r < a. \end{cases} \tag{2.29}$$

Eqs. (2.25) and (2.26) do not define the stresses uniquely. Let us consider the rigid axisymmetric punch given in Fig. 1 with a radius $a > 0$. Assume that the indentation is monotonically increased until $w(r) = \delta > 0$, at a sufficiently slow rate to permit a quasi-static treatment. Note that if $w(s) = \delta$, $0 \leqslant s \leqslant r$, then

$$\frac{\mathrm{d}}{\mathrm{d}r}\left[\int_0^r \frac{sw(s)}{(r^2-s^2)^{1/2}}\,\mathrm{d}s\right] = \delta.$$

Writing the stresses and displacements non-dimensionally as

$$\begin{cases} \sigma(r) = -\dfrac{\mu}{1-v}\dfrac{\delta}{a}p(x), \\ \tau(r) = -\dfrac{\mu}{1-v}\dfrac{\delta}{a}q(x), \\ w(r) = \delta, \\ u(r) = \delta U(x), \end{cases} \tag{2.30}$$

where $x = r/a$, Eqs. (2.25) and (2.26) may be written as, using the boundary conditions (2.29),

$$\int_x^1 \frac{tp(t)}{\sqrt{t^2-x^2}}\,\mathrm{d}t - \gamma\left\{\int_0^1 q(t)\,\mathrm{d}t - x\int_0^x \frac{q(t)}{\sqrt{x^2-t^2}}\,\mathrm{d}t\right\} = 1, \tag{2.31}$$

$$x\int_x^1 \frac{q(t)}{\sqrt{t^2-x^2}}\,\mathrm{d}t - \gamma\int_0^x \frac{tp(t)}{\sqrt{x^2-t^2}}\,\mathrm{d}t = U^*(x), \tag{2.32}$$

for $x \in]0,1[$ and U^* is defined by

$$U^*(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left[x\int_0^x \frac{U(s)}{(x^2-s^2)^{1/2}}\,\mathrm{d}s\right].$$

We look for a solution in which no slip takes place over a central circle $(0, c)$ $(c < 1)$. The frictional force required to resist slip at points within the circle cannot then exceed the limiting value, while in the outer annulus $[c, 1]$, as the normal force is slowly increased, slip takes place inwards, with the limiting friction acting outwards, so that

$$U(x) = 0 \quad \text{and} \quad \mu p(x) - q(x) > 0, \quad x \in [0, c), \tag{2.33}$$

$$U(x) < 0 \quad \text{and} \quad \mu p(x) - q(x) = 0, \quad x \in [c, 1]. \tag{2.34}$$

Henceforth we shall refer to this as the coupled Abel's integral equations free point problem.

3. Coupled Volterra integral equations and product integration

In the previous section we derived a characterisation of the Hertz contact problem with finite friction in terms of the coupled Abel's equations (2.31)–(2.32). In this section we denote the function U by u. The system can be written in two equivalent forms. The first one was employed by Spence [15], the second is more convenient for analysis as its discretisation yields matrices whose structure is essentially Toeplitz.

The system (2.31)–(2.32) can be rewritten as follows:

$$\int_x^1 \frac{tp(t)}{\sqrt{t^2-x^2}}\,\mathrm{d}t - \gamma\left\{\int_0^1 q(t)\,\mathrm{d}t - x\int_0^x \frac{q(t)}{\sqrt{x^2-t^2}}\,\mathrm{d}t\right\} = 1, \tag{3.1}$$

$$\gamma\int_0^x \frac{tp(t)}{\sqrt{x^2-t^2}}\,\mathrm{d}t - x\int_x^1 \frac{q(t)}{\sqrt{t^2-x^2}}\,\mathrm{d}t + \frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\left[\int_0^x \frac{t^2u(t)}{\sqrt{x^2-t^2}}\,\mathrm{d}t\right] = 0 \tag{3.2}$$

subject to

$$u(x) = 0 \quad \text{and} \quad z(x) = q(x) - \mu p(x) < 0, \quad x \in [0, c), \tag{3.3}$$

$$u(x) < 0 \quad \text{and} \quad z(x) = q(x) - \mu p(x) = 0, \quad x \in [c, 1]. \tag{3.4}$$

Then using the change of variable

$$\begin{cases} s = t^2, \\ X = x^2, \end{cases} \tag{3.5}$$

gives

$$\int_X^1 \frac{P(s)}{2\sqrt{s-X}}\,\mathrm{d}s - \gamma\left\{\int_0^1 \frac{1}{2}Q(s)\,\mathrm{d}s - \sqrt{X}\int_0^X \frac{Q(s)}{2\sqrt{X-S}}\,\mathrm{d}s\right\} = 1, \tag{3.6}$$

$$\gamma\int_0^X \frac{P(s)}{2\sqrt{X-s}}\,\mathrm{d}s - \sqrt{X}\int_X^1 \frac{Q(s)}{2\sqrt{s-X}}\,\mathrm{d}s + \frac{\mathrm{d}}{\mathrm{d}X}\left[\int_0^X \frac{U(s)}{\sqrt{X-s}}\,\mathrm{d}s\right] = 0, \tag{3.7}$$

where, for $x \in]0, 1[$,

$$\begin{cases} P(x) = p(\sqrt{x}), \\ Q(x) = \dfrac{q(\sqrt{x})}{\sqrt{x}}, \\ U(x) = \sqrt{x}u(\sqrt{x}). \end{cases} \tag{3.8}$$

The complementarity conditions (3.3)–(3.4) transform to give

$$U(X) = 0 \quad \text{and} \quad Z(X) = \sqrt{X}Q(X) - \mu P(X) < 0, \quad X \in [0, c^2), \tag{3.9}$$

$$U(X) < 0 \quad \text{and} \quad Z(X) = \sqrt{X}Q(X) - \mu P(X) = 0, \quad X \in [c^2, 1]. \tag{3.10}$$

Discretise [0, 1] with the mesh points $s_j = jh(j = 0, \ldots, N)$ where $h = 1/N$ and $N > 0$. Let

$$\begin{cases} X_i = \frac{1}{2}(s_{i-1} + s_i), \\ P_i = P(X_i), \quad Q_i = Q(X_i), \quad U_i = U(X_i), \end{cases}$$

for $1 \leqslant i \leqslant N$ and rewrite (3.6)–(3.7) in product integration form. The following two integrals for $X \in (s_{i-1}, s_i]$ are discretised using the summations, $1 \leqslant i \leqslant N$:

$$\int_0^X f(X, s)\,\mathrm{d}s = \sum_{j=1}^{i-1}\int_{s_{j-1}}^{s_j} f(X, s)\,\mathrm{d}s + \int_{s_{i-1}}^X f(X, s)\,\mathrm{d}s,$$

$$\int_X^1 f(X, s)\,\mathrm{d}s = \int_X^{s_i} f(X, s)\,\mathrm{d}s + \sum_{j=i+1}^{N}\int_{s_{j-1}}^{s_j} f(X, s)\,\mathrm{d}s.$$

Thus we obtain

$$\int_{X_i}^{s_i} \frac{1}{2\sqrt{s-X_i}}\,\mathrm{d}s\,P_i + \sum_{j=i+1}^{N}\int_{s_{j-1}}^{s_j} \frac{1}{2\sqrt{s-X_i}}\,\mathrm{d}s\,P_j - \gamma\left\{\sum_{j=1}^{N}\int_{s_{j-1}}^{s_j}\frac{1}{2}\,\mathrm{d}s\,Q_j\right.$$
$$\left. -\sqrt{X_i}\left[\sum_{j=1}^{i-1}\int_{s_{j-1}}^{s_j}\frac{1}{2\sqrt{X_i-s}}\,\mathrm{d}s\,Q_j + \int_{s_{i-1}}^{X_i}\frac{1}{2\sqrt{X_i-s}}\,\mathrm{d}s\,Q_i\right]\right\} = 1 \tag{3.11}$$

and

$$\sqrt{X_i}\left[\int_{X_i}^{s_i}\frac{1}{2\sqrt{s-X_i}}\,\mathrm{d}s\,Q_i + \sum_{j=i+1}^{N}\int_{s_{j-1}}^{s_j}\frac{1}{2\sqrt{s-X_i}}\,\mathrm{d}s\,Q_j\right]$$
$$-\gamma\left[\sum_{j=1}^{i-1}\int_{s_{j-1}}^{s_j}\frac{1}{2\sqrt{X_i-s}}\,\mathrm{d}s\,P_j + \int_{s_{i-1}}^{X_i}\frac{1}{2\sqrt{X_i-s}}\,\mathrm{d}s\,P_i\right] = U_i^*, \tag{3.12}$$

where

$$U_i^* = \frac{2}{h}\left(\sum_{j=1}^{i-1}\int_{s_{j-1}}^{s_j}\frac{1}{2\sqrt{X_i-s}}\,\mathrm{d}s U_j + \int_{s_{i-1}}^{X_i}\frac{1}{2\sqrt{X_i-s}}\,\mathrm{d}s\,U_i\right.$$
$$\left.-\left[\sum_{j=1}^{i-2}\int_{s_{j-1}}^{s_j}\frac{1}{2\sqrt{X_{i-1}-s}}\,\mathrm{d}s U_j + \int_{s_{i-2}}^{X_{i-1}}\frac{1}{2\sqrt{X_{i-1}-s}}\,\mathrm{d}s U_{i-1}\right]\right) \tag{3.13}$$

subject to

$$U_i = 0 \quad \text{and} \quad Z_i < 0, \quad 1\leqslant i < v, \tag{3.14}$$

$$U_i < 0 \quad \text{and} \quad Z_i = 0, \quad v\leqslant i\leqslant N \tag{3.15}$$

for some (unknown) integer v. Note that interpreting $i-1$ when $i=1$ is not a problem if the mesh is taken fine enough since we are assuming the existence of some point c ($\neq 0$) where, for $x<c$, no slippage takes place, that is $u(x)\equiv 0$ in that interval and consequently the integrals in (3.13) are identically zero. The existence of a unique point c has been demonstrated by Spence [15] by writing the problem as a Fredholm integral equation for the function $\varphi(x) = p(x) - q(x)/\mu$.

The integrals are evaluated as follows:

$$I_{ij} = \int_{s_{j-1}}^{s_j}\frac{1}{2\sqrt{|s-X_i|}}\,\mathrm{d}s = \begin{cases} \int_{s_{j-1}}^{s_j}\dfrac{1}{2\sqrt{s-X_i}}\,\mathrm{d}s, & 1\leqslant i<j\leqslant N,\\ \int_{s_{j-1}}^{s_j}\dfrac{1}{2\sqrt{X_i-s}}\,\mathrm{d}s, & 1\leqslant j<i\leqslant N\end{cases}$$

and, for $1\leqslant i\leqslant N$,

$$\begin{cases} I_{ii}^- = \int_{s_{i-1}}^{X_i}\dfrac{1}{2\sqrt{X_i-s}}\,\mathrm{d}s,\\ I_{ii}^+ = \int_{X_i}^{s_i}\dfrac{1}{2\sqrt{s-X_i}}\,\mathrm{d}s.\end{cases}$$

We obtain for $i\neq j$,

$$I_{ij} = \left|\sqrt{|s_j-X_i|} - \sqrt{|s_{j-1}-X_i|}\right|$$

and so

$$\begin{cases} I_{ij} = \sqrt{h}\left|\sqrt{|j-i+\frac{1}{2}|} - \sqrt{|j-i-\frac{1}{2}|}\right| = \sqrt{h}|\alpha^{+}_{|i-j|} - \alpha^{-}_{|i-j|}| = \sqrt{h}\alpha_{|i-j|},\ i \neq j, \\ I^{-}_{ii} = \sqrt{h}\alpha^{-}_{0} = \sqrt{h}\sqrt{\frac{1}{2}}, \\ I^{+}_{ii} = \sqrt{h}\alpha^{+}_{0} = \sqrt{h}\sqrt{\frac{1}{2}}. \end{cases} \tag{3.16}$$

We can therefore define the $N \times N$ matrix $\mathbf{A}$ by its elements $\alpha_{|i-j|}$ given in (3.16). In this formulation, the matrix $\mathbf{A}$ characterising discretised equations is both symmetric and Toeplitz. Its upper and lower triangular parts are, respectively, defined from (3.16) by

$$\mathbf{A}^{\mathrm{U}}_{ij} = \begin{cases} \alpha_{j-i}, & i < j, \\ \alpha^{+}_{0}, & i = j, \\ 0, & i > j, \end{cases} \quad \text{and} \quad \mathbf{A}^{\mathrm{L}}_{ij} = \mathbf{A}^{\mathrm{U}}_{ji}, \quad 1 \leqslant i, \quad j \leqslant N.$$

Define the following matrices:

$$\mathbf{\Gamma}_{ij} = \tfrac{1}{2}, \quad 1 \leqslant i, \quad j \leqslant N,$$

$$\mathbf{D}_{ij} = \begin{cases} \sqrt{i - \frac{1}{2}} & \text{if } i = j, \\ 0 & \text{otherwise,} \end{cases} \quad 1 \leqslant i, \quad j \leqslant N,$$

$$\mathbf{\Delta} = \begin{pmatrix} 1 & 0 & & \cdots & & 0 \\ -1 & 1 & 0 & \cdots & 0 & \\ 0 & -1 & 1 & & \vdots & \vdots \\ \vdots & \vdots & & & 0 & \\ 0 & 0 & \ldots & -1 & 1 & 0 \\ 0 & 0 & \ldots & 0 & -1 & 1 \end{pmatrix}_{N \times N},$$

$$\mathbf{1} = (1, 1, \ldots, 1)^{\mathrm{T}}_{N \times 1}.$$

Note $(1/h\mathbf{\Delta})$ is sometimes known as the differentiation matrix (e.g., [10]).

The matrix formulation of (3.11)–(3.12) is therefore given by

$$\sqrt{h}\mathbf{A}^{\mathrm{U}}\mathbf{P} - \gamma h(\mathbf{\Gamma}\mathbf{Q} - \mathbf{D}\mathbf{A}^{\mathrm{L}}\mathbf{Q}) = \mathbf{1}, \tag{3.17}$$

$$-\gamma\sqrt{h}\mathbf{A}^{\mathrm{L}}\mathbf{P} + h\mathbf{D}\mathbf{A}^{\mathrm{U}}\mathbf{Q} - \frac{2}{\sqrt{h}}\mathbf{\Delta}\mathbf{A}^{\mathrm{L}}\mathbf{U} = \mathbf{0} \tag{3.18}$$

and

$$-\mathbf{U}^{\mathrm{T}}(\mu\mathbf{P} - \sqrt{h}\mathbf{D}\mathbf{Q}) = 0, \tag{3.19}$$

$$-\mathbf{U} \geqslant \mathbf{0} \quad \text{and} \quad \mu\mathbf{P} - \sqrt{h}\mathbf{D}\mathbf{Q} \geqslant \mathbf{0}, \tag{3.20}$$

where the inequalities between vectors are satisfied for all their components.

4. Formulation as a mixed LCP

The LCP consists of finding a vector in a finite-dimensional real vector space that satisfies a certain system of inequalities. Specifically, given a vector $\mathbf{q} \in \mathbb{R}^n$ and a matrix $\mathbf{M} \in \mathbb{R}^{n \times m}$, the LCP is to find a vector $\mathbf{z} \in \mathbb{R}^n$ such that

$$\mathbf{z} \geqslant \mathbf{0}, \tag{4.1}$$

$$\mathbf{q} + \mathbf{M}\mathbf{z} \geqslant \mathbf{0}, \tag{4.2}$$

$$\mathbf{z}^{\mathrm{T}}(\mathbf{q} + \mathbf{M}\mathbf{z}) = 0 \tag{4.3}$$

or to show that no such vector $\mathbf{z}$ exists. Henceforth, the LCP given by (4.1)–(4.3) will be referred to as LCP($\mathbf{q}$, $\mathbf{M}$). The mixed LCP, on the other hand, involves an additional system of linear equations.

Let $\mathscr{A}$ and $\mathscr{B}$ be real square matrices of orders n and m, respectively. Let $\mathscr{C} \in \mathbb{R}^{n\times m}$, $\mathscr{D} \in \mathbb{R}^{m\times n}$, $\mathbf{a} \in \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^m$ be given. The mixed LCP is to find the vectors

$$\mathbf{v} \in \mathbb{R}^n, \quad \mathbf{w} \in \mathbb{R}^m,$$

such that

$$\mathbf{a} + \mathscr{A}\mathbf{v} + \mathscr{C}\mathbf{w} = \mathbf{0}, \tag{4.4}$$

$$\mathbf{b} + \mathscr{D}\mathbf{v} + \mathscr{B}\mathbf{w} \geqslant \mathbf{0}, \tag{4.5}$$

$$\mathbf{w} \geqslant \mathbf{0}, \tag{4.6}$$

$$\mathbf{w}^{\mathrm{T}}(\mathbf{b} + \mathscr{D}\mathbf{v} + \mathscr{B}\mathbf{w}) = \mathbf{0}. \tag{4.7}$$

Thus, as stated above, the mixed LCP is simply a LCP combined with a system of linear equations.

For (4.4)–(4.7), if the matrix $\mathscr{A}$ is non-singular, we may solve for the vector $\mathbf{v}$, obtaining

$$\mathbf{v} = -\mathscr{A}^{-1}(\mathbf{a} + \mathscr{C}\mathbf{w}).$$

Note that the variable $\mathbf{v}$ is not restricted to be non-negative. By eliminating $\mathbf{v}$ in the remaining conditions of the problem (4.4)–(4.7), we can convert this mixed LCP into the standard LCP($\mathbf{q}$, $\mathbf{M}$) with

$$\mathbf{q} = \mathbf{b} - \mathscr{D}\mathscr{A}^{-1}\mathbf{a}, \quad \mathbf{M} = \mathscr{B} - \mathscr{D}\mathscr{A}^{-1}\mathscr{C}.$$

Although this may be useful from a theoretical point of view, it is not always advisable numerically.

Let us rewrite (3.17)–(3.20) in the form

$$\begin{pmatrix} \sqrt{h}\mathbf{A}^{\mathrm{U}} & -\gamma h\mathbf{G} & \mathbf{0} \\ -\gamma\sqrt{h}\mathbf{A}^{\mathrm{L}} & h\mathbf{D}\mathbf{A}^{\mathrm{U}} & \frac{2}{\sqrt{h}}\mathbf{\Delta}\mathbf{A}^{\mathrm{L}} \\ \mu\mathbf{I} & -\sqrt{h}\mathbf{D} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{p} \\ \mathbf{q} \\ \mathbf{u}^* \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{r} \end{pmatrix} + \begin{pmatrix} \mathbf{1} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} \tag{4.8}$$

subject to

$$\mathbf{u}^* \geqslant \mathbf{0}, \quad \mathbf{r} \geqslant \mathbf{0}, \quad \mathbf{r}^{\mathrm{T}}\mathbf{u}^* = 0, \tag{4.9}$$

where $\mathbf{G} = (\mathbf{\Gamma} - \mathbf{D}\mathbf{A}^{\mathrm{L}})$, $\mathbf{I} \in \mathbb{R}^{N\times N}$ is the identity matrix, $\mathbf{u}^* = -\mathbf{u}$ and we have replaced $\mathbf{P}$ by $\mathbf{p}$, $\mathbf{Q}$ by $\mathbf{q}$, $\mathbf{U}$ by $\mathbf{u}$ and $\mathbf{r} = \mu\mathbf{p} - \sqrt{h}\,\mathbf{D}\mathbf{q}$.

Appropriate identification of the matrices and vectors shows that the finite difference discretisation of the Abel integral equations characterisation of the Hertz contact problem with finite friction is indeed a mixed LCP.

5. Numerical methods

Billups et al. [3] carried out an excellent comparison of large scale nonlinear mixed complementarity problem (MCP) solvers. They showed that two numerical methods are particularly robust and efficient.

The first method was developed by Dirkse and Ferris [8] and was based on Newton's method for solving non-smooth equations. This method provides a convergent iterative scheme thus reducing the nonlinear problem to a sequence of mixed LCPs.

The second scheme was developed by Chen and Mangasarian [6]. The formulation of the MCP uses the plus function, $(x)_+ = \max(0, x)$ where $x \in \mathbb{R}$, which is then approximated by a smooth parametric function. This leads to a class of smooth parametric nonlinear equations that can be solved using a classical Newton-based algorithm.

The system (4.8)–(4.9) is equivalent to the following system: find the vector $(\mathbf{p}, \mathbf{q}, \mathbf{u}^*) \in \mathbb{R}^{3N}$ such that

$$\begin{aligned}
&\sqrt{h}\mathbf{A}^{\mathrm{U}}\mathbf{p} - \gamma h \mathbf{G}\mathbf{q} = \mathbf{0},\\
&-\gamma\sqrt{h}\mathbf{A}^{\mathrm{L}}\mathbf{p} + h\mathbf{D}\mathbf{A}^{\mathrm{U}}\mathbf{q} + \frac{2}{\sqrt{h}}\boldsymbol{\Delta}\mathbf{A}^{\mathrm{L}}\mathbf{u}^* = \mathbf{0},\\
&\mathbf{u}^* - \max(\mathbf{0}, \mathbf{u}^* - (\mu\mathbf{p} - \sqrt{h}\mathbf{D}\mathbf{q})) = \mathbf{0},
\end{aligned} \tag{5.1}$$

where the plus function is applied to each component of the third equation.

We shall denote this system by

$$\mathbf{M}(\mathbf{p}, \mathbf{q}, \mathbf{u}^*) = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}.$$

A smooth approximation of the system (5.1) can be defined using the approximation of the plus function in the following sense. The derivative of the plus function is the Heaviside step function: for all $x \in \mathbb{R}\backslash\{0\}$,

$$(x)'_+ = H(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x < 0 \end{cases}$$

and the derivative of $H(x)$ (in the sense of distribution theory) is the Dirac measure $\delta(x)$. Following Chen and Mangasarian [6] we introduce the smooth parametric approximation of $\delta(x)$ by

$$p(x, \beta) = x + \beta \log(1 + \mathrm{e}^{-x/\beta}), \quad \beta > 0. \tag{5.2}$$

Integration of p provides an approximation of the plus function.

By using $p(x, \beta)$ instead of the plus function, the MCP (5.1) can be approximated as follows: find the vector $(\mathbf{p}, \mathbf{q}, \mathbf{u}^*) \in \mathbb{R}^{3N}$ such that

$$\hat{\mathbf{M}}(\mathbf{p}, \mathbf{q}, \mathbf{u}^*) = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \tag{5.3}$$

where the third component, $M_3(\mathbf{p}, \mathbf{q}, \mathbf{u}^*)$, has been replaced by

$$u_i^* - p(u_i^* - (\mu\mathbf{p} - \sqrt{h}\mathbf{D}\mathbf{q})_i, \beta) = 0, \quad 1 \leqslant i \leqslant N. \tag{5.4}$$

System (5.3) is now continuously differentiable. The strategy is to compute the solution of (5.3) by minimising $f(\mathbf{p}, \mathbf{q}, \mathbf{u}^*)$ defined by

$$f(\mathbf{p}, \mathbf{q}, \mathbf{u}^*) = \tfrac{1}{2}\hat{\mathbf{M}}(\mathbf{p}, \mathbf{q}, \mathbf{u}^*)^{\mathrm{T}}\hat{\mathbf{M}}(\mathbf{p}, \mathbf{q}, \mathbf{u}). \tag{5.5}$$

The choice of the parameter β in (5.2) is dependent on the evaluation point $(\mathbf{p}, \mathbf{q}, \mathbf{u}^*)$ of the function $\hat{\mathbf{M}}$:

$$\frac{1}{\beta} = \alpha(\mathbf{p}, \mathbf{q}, \mathbf{u}^*) = \begin{cases} \dfrac{\sqrt{N}}{\|\hat{\mathbf{M}}(\mathbf{p}, \mathbf{q}, \mathbf{u}^*)\|_2} & \text{if } \|\hat{\mathbf{M}}(\mathbf{p}, \mathbf{q}, \mathbf{u}^*)\|_2 \leqslant \sqrt{N}, \\ \sqrt{\dfrac{\sqrt{N}}{\|\hat{\mathbf{M}}(\mathbf{p}, \mathbf{q}, \mathbf{u}^*)\|_2}} & \text{otherwise,} \end{cases} \tag{5.6}$$

where $\|\cdot\|_2$ is the usual Euclidean norm in $\mathbb{R}^{3N}$.

The resulting algorithm, due to Chen and Mangasarian [6], is given in the Appendix. Conditions for its convergence, as it pertains to the method of this paper, are given in the following theorem.

Theorem 1. *Let us consider the mixed LCP* (5.1).
If the matrix

$$\mathbf{M}_{\mathrm{D}} = \begin{pmatrix} \sqrt{h}\mathbf{A}^{\mathrm{U}} & -\gamma h\mathbf{G} & \mathbf{0} \\ -\gamma\sqrt{h}\mathbf{A}^{\mathrm{L}} & h\mathbf{D}\mathbf{A}^{\mathrm{U}} & \frac{2}{\sqrt{h}}\boldsymbol{\Delta}\mathbf{A}^{\mathrm{L}} \\ \mu\mathbf{I} & -\sqrt{h}\mathbf{D} & \mathscr{D} \end{pmatrix} \tag{5.7}$$

is invertible for all positive diagonal matrices $\mathscr{D} \in \mathbb{R}^{N\times N}$, *then*

(1) *The sequence* $(\mathbf{p}^k, \mathbf{q}^k, \mathbf{u}^{*^k})_{k\geqslant 0}$ *defined in Algorithm* 1 (*see Appendix*) *exists.*
(2) *Any accumulation point of the above sequence is an* ε*-accurate solution of the MCP* (5.1).
(3) *If an accumulation point exists, the whole sequence* $(\mathbf{p}^k, \mathbf{q}^k, \mathbf{u}^{*^k})_{k\geqslant 0}$ *converges to an* ε*-accurate solution quadratically.*
(4) *If in addition, the level set*

$$\left\{(\mathbf{p}, \mathbf{q}, \mathbf{u}^*) : \|\hat{\mathbf{M}}(\mathbf{p}, \mathbf{q}, \mathbf{u}^*)\|_2 \leqslant \|\hat{\mathbf{M}}(\mathbf{p}^0, \mathbf{q}^0, \mathbf{u}^{*0})\|_2 + \frac{\eta}{\eta-1}\frac{2\sqrt{N}\log(2)}{\alpha^0}\right\} \tag{5.8}$$

is compact, the sequence $(\mathbf{p}^k, \mathbf{q}^k, \mathbf{u}^{*^k})_{k\geqslant 0}$ *converges to an* ε*-accurate solution at a quadratic rate.* (*Note* $\alpha^0 = \alpha(\mathbf{p}^0, \mathbf{q}^0, \mathbf{u}^{*^0})$ *and* η *is a given tolerance.*)

Proof. See Theorem 4.3 in [6]. □

If (3.1)–(3.2) is discretised directly, rather than using the change of variables (3.5), the resulting discretisation can be shown to be a mixed LCP and Theorem 1 may be applied. In this case convergence of Algorithm 1 depends upon the following matrix being invertible:

$$\tilde{\mathbf{M}}_{\mathrm{D}} = \begin{pmatrix} h\tilde{\mathbf{A}}^{\mathrm{U}} & -\gamma h\tilde{\mathbf{G}} & \mathbf{0} \\ -\gamma h\tilde{\mathbf{A}}^{\mathrm{L}} & h^2\tilde{\mathbf{D}}\tilde{\mathbf{A}}^{\mathrm{U}} & \tilde{\mathbf{D}}^{-1}\boldsymbol{\Delta}\tilde{\mathbf{A}}^{\mathrm{L}}\tilde{\mathbf{D}} \\ \mu\mathbf{I} & -h\tilde{\mathbf{D}} & \mathscr{D} \end{pmatrix}, \tag{5.9}$$

where $\tilde{\boldsymbol{\Gamma}}_{ij} = j - \frac{1}{2}$, $1\leqslant i, j\leqslant N$, $\tilde{\mathbf{D}} = \mathbf{D}^2$, $\tilde{\mathbf{G}} = \tilde{\boldsymbol{\Gamma}} - \tilde{\mathbf{D}}\tilde{\mathbf{A}}^{\mathrm{L}}$, and

$$\tilde{\mathbf{A}}^{\mathrm{U}}_{ij} = \begin{cases} \tilde{\alpha}_{ij}, & i<j, \\ \tilde{\alpha}^{+}_{ii}, & i=j, \\ 0, & i>j, \end{cases} \quad \text{and} \quad \tilde{\mathbf{A}}^{\mathrm{L}} = \begin{cases} \tilde{\alpha}_{ij}, & i>j, \\ \tilde{\alpha}^{-}_{ii}, & i=j, \\ 0, & i\leqslant j \end{cases}$$

with

$$\begin{aligned} \tilde{\alpha}_{ij} &= |\tilde{\alpha}^{+}_{ij} - \tilde{\alpha}^{-}_{ij}| = \left|\sqrt{|j^2 - \left(i-\tfrac{1}{2}\right)^2|} - \sqrt{|(j-1)^2 - (i-\tfrac{1}{2})^2|}\right|, \\ \tilde{\alpha}^{-}_{ii} &= \sqrt{|(i-1)^2 - (i-\tfrac{1}{2})^2|}, \\ \tilde{\alpha}^{+}_{ii} &= \sqrt{|i^2 - (i-\tfrac{1}{2})^2|}. \end{aligned} \tag{5.10}$$

6. Convergence of the Chen and Mangasarian algorithm

Theorem 1 is satisfied if the matrix $\mathbf{M}_{\mathrm{D}}$ defined in (5.7) is invertible. We write

$$\mathbf{M}_{\mathrm{D}} = \begin{pmatrix} \sqrt{h}\mathbf{U} & -\gamma h\mathbf{G} & \mathbf{0} \\ -\gamma\sqrt{h}\mathbf{U}^{\mathrm{T}} & h\mathbf{D}\mathbf{U} & \frac{2}{\sqrt{h}}\boldsymbol{\Delta}\mathbf{U}^{\mathrm{T}} \\ \mu\mathbf{I} & -\sqrt{h}\mathbf{D} & \mathscr{D} \end{pmatrix},$$

where the various submatrices and parameters are as in (5.7) but to avoid excessive use of subscripts we let $\mathbf{U} = \mathbf{A}^{\mathrm{U}}$ (hence $\mathbf{U}^{\mathrm{T}} = \mathbf{A}^{\mathrm{L}}$). We let the diagonal elements of the (positive diagonal) matrix $\mathscr{D}$ be $\lambda_1, \ldots, \lambda_N$. Note that if the blocks of $\mathbf{M}_{\mathrm{D}}$ are $N \times N$ then $h = 1/N$.

We shall prove that this matrix is invertible for various values of the parameters γ and μ, provided the following lemma holds.

Lemma 1. *If $U \in \mathbb{R}^{N\times N}$ is the upper triangular Toeplitz matrix whose first row is*

$$u_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & \sqrt{3}-1 & \sqrt{5}-\sqrt{3} & \cdots & \sqrt{2N+1}-\sqrt{2N-1} \end{pmatrix}$$

then for all N, $\|U^{-1}\|_1 < 3$.

We also need the block version of the Sherman–Morrison formula.

Theorem 2. *Let $A \in \mathbb{R}^{N\times N}$ be non-singular and $B, C \in \mathbb{R}^{N\times K}$. Then $A + BC^{\mathrm{T}}$ is non-singular if and only if $D = (I + C^{\mathrm{T}}A^{-1}B)$ is non-singular, in which case*

$$(A + BC^{\mathrm{T}})^{-1} = A^{-1} - A^{-1}BD^{-1}C^{\mathrm{T}}A^{-1}.$$

Proof. See [9, p. 258], for example. □

Theorem 3. *Define $\mathbf{M}_{\mathrm{D}}$ as above. Then $\mathbf{M}_{\mathrm{D}}$ is non-singular if*

$$\tfrac{3}{2}\{\gamma\mu + (1+\gamma^2)\lambda_{\max}\}\sqrt{N} < 1.$$

First, we simplify the matrix a little. Let $\mathbf{D}_1 = \mathrm{diag}(\sqrt{h}\mathbf{I}, \sqrt{h}\mathbf{I}, \mathbf{I})$ and $\mathbf{D}_2 = \mathrm{diag}(\mathbf{I}, \sqrt{h}\mathbf{I}, \mathbf{I})$. We can write

$$\mathbf{M}_{\mathrm{D}} = \mathbf{D}_1 \begin{pmatrix} \mathbf{U} & -\gamma\mathbf{G} & \mathbf{0} \\ \mathbf{0} & \mathbf{DU} & \dfrac{2}{h}\mathbf{U}^{\mathrm{T}}\boldsymbol{\Delta} \\ \mathbf{B} & -\mathbf{D} & \mathscr{D} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \gamma h\boldsymbol{\Delta}^{-1} & \mathbf{0} & \mathbf{I} \end{pmatrix} \mathbf{D}_2,$$

where $\mathbf{B} = \mu\mathbf{I} + \gamma h\mathscr{D}\boldsymbol{\Delta}^{-1}$. Note we have used the fact that $\boldsymbol{\Delta}\mathbf{U}^{\mathrm{T}} = \mathbf{U}^{\mathrm{T}}\boldsymbol{\Delta}$ (triangular Toeplitz matrices commute).

Any questions about the invertibility of $\mathbf{M}_{\mathrm{D}}$ can be answered by considering

$$\mathbf{X} = \begin{pmatrix} \mathbf{U} & -\gamma\mathbf{G} & \mathbf{0} \\ \mathbf{0} & \mathbf{DU} & \tfrac{2}{h}\mathbf{U}^{\mathrm{T}}\boldsymbol{\Delta} \\ \mathbf{B} & -\mathbf{D} & \mathscr{D} \end{pmatrix}.$$

Clearly, the upper triangular part of $\mathbf{X}$ is non-singular. For now we express this triangular inverse in block form as

$$\begin{pmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \mathbf{X}_3 \\ \mathbf{0} & \mathbf{X}_4 & \mathbf{X}_5 \\ \mathbf{0} & \mathbf{0} & \mathbf{X}_6 \end{pmatrix}.$$

Applying the Sherman–Morrison formula, we find that $\mathbf{X}$ is non-singular if and only if the matrix $\mathbf{Z} = \mathbf{I} + \mathbf{B}\mathbf{X}_3 - \mathbf{D}\mathbf{X}_5$ is, too. Computing $\mathbf{X}_3$ and $\mathbf{X}_5$ explicitly gives

$$\begin{aligned} \mathbf{Z} &= \mathbf{I} - \frac{\gamma}{h}\mathbf{B}\mathbf{U}^{-1}\mathbf{G}\mathbf{U}^{-1}\mathbf{D}^{-1}\mathbf{U}^{\mathrm{T}}\boldsymbol{\Delta}\mathscr{D}^{-1} + \frac{2}{h}\mathbf{D}\mathbf{U}^{-1}\mathbf{D}^{-1}\mathbf{U}^{\mathrm{T}}\boldsymbol{\Delta}\mathscr{D}^{-1} \\ &= \frac{2}{h}\left(\mathbf{I} - \frac{\gamma}{2}\mathbf{B}\mathbf{U}^{-1}\mathbf{G}\mathbf{D}^{-1} + \frac{h}{2}\mathscr{D}\boldsymbol{\Delta}^{-1}\mathbf{U}^{-\mathrm{T}}\mathbf{D}\mathbf{U}\mathbf{D}^{-1}\right)\mathbf{D}\mathbf{U}^{-1}\mathbf{D}^{-1}\mathbf{U}^{\mathrm{T}}\boldsymbol{\Delta}\mathscr{D}^{-1}. \end{aligned}$$

We conclude that $\mathbf{M}_{\mathrm{D}}$ is non-singular if and only if this final bracketed expression is non-singular. To get a sufficient condition for non-singularity we simply ensure that the norm of the sum of the final two terms in the brackets is bounded above by 1.

Using the structure of $\mathbf{D}$ and $\mathbf{G}$, it is straightforward to show that $\|\mathbf{GD}^{-1}\|_1 < \sqrt{N}$ and $\|\mathbf{DUD}^{-1}\|_1 < \sqrt{N}$ for all N. A bound is then established by noting that $\|h\mathbf{\Delta}^{-1}\|_1 < 1$ and applying Lemma 1, since $\mathbf{U} = U$.

The dependence on $\sqrt{N}$ is rather restrictive. To do better we need information on the size of the elements of U^{-1} along with its norm. Numerical evidence suggests that the elements on the kth diagonal of U^{-1} converge quickly to $-1/(\pi k^{3/2})$. We can then show that $\|\mathbf{U}^{-1}\mathbf{GD}^{-1}\|_1$ and $\|\mathbf{U}^{-\mathrm{T}}\mathbf{DUD}^{-1}\|_1$ are bounded above by $\log N$ for all N. Our condition for non-singularity then becomes

$$\gamma\mu + (1+\gamma^2)\lambda_{\max} < \frac{2}{\log N}.$$

Alternatively, we can use the bound $\|h\mathscr{D}\mathbf{\Delta}^{-1}\|_1 < \sum \lambda_i/N$ to get the condition

$$\frac{\log N}{N}\left(\gamma\mu + (1+\gamma^2)\sum_{i=1}^{N}\lambda_i\right) < 2.$$

If we know that $\sum_{i=1}^{N}\lambda_i$ grows sufficiently slowly then we can conclude that $\mathbf{M}_{\mathrm{D}}$ will always be non-singular for sufficiently large N (for a rigorous bound, we replace $(\log N)/N$ with $3N^{-1/2}$).

However, Lemma 1 remains unproven. This is perhaps not surprising as this is a special case of a classical open problem in the theory of the convergence of discretisation methods for Abel's equations of some 30 years standing. For a more general exposition of this question the reader is referred to Brunner [4,5].

7. Numerical results

To validate the method of this paper, a comparison is performed between the mixed LCP (4.8)–(4.9) and Spence's method. It should be recalled that Spence does not compute the solution directly, but rather guesses the value of c, the point where adhesion terminates and slippage commences, and iterates until he obtains a value of c which gives the correct value of the friction coefficient μ. He achieves this using a *regula falsi* iterative method. In contrast, the method proposed in this paper computes $(p(x), q(x), u(x))$ and c directly through a mixed LCP formulation.

Table 1 shows that the results obtained by the mixed LCP formulation are in general agreement with the rather limited results provided by Spence. These have been computed for various values of the parameters v and μ. Note that (MLCP1) refers to the problem where the transformation of variables (3.5) have been employed, whereas (MLCP2) refers to the untransformed problem. The difference between the results of (MLCP1) and (MLCP2) are due to the scaling effect of the transformation (3.5): the mesh spacing becomes larger near $x = 0$ causing imprecision and smaller near $x = 1$ resulting in greater precision.

Fig. 2 shows that the complementarity condition is respected: the ratio $q(x)/\mu p(x)$ is also the function that Spence uses for fitting the constraint (3.3)–(3.4) in his algorithm. Figs. 3 and 4 represent, respectively, the profile of the normal stress p and the tangential shear stress q computed by (MLCP1) and (MLCP2). The difference is small and due entirely

Table 1
Comparison of the values of μ and c found by Spence's method [15] and mixed LCP formulation

	μ	0.1387	0.1801	0.2063	0.2843	0.2986	0.4013	0.4862
$v = 0$	c (Spence)			0.24		0.5	0.7	0.8
	c (MLCP1)	0.075	0.175	0.2417	0.475	0.5083	0.7083	0.8083
	c (MLCP2)	0.0913	0.2041	0.2739	0.4743	0.5083	0.7072	0.8010
$v = 0.25$	c (Spence)	0.3	0.5		0.8			
	c (MLCP1)	0.3083	0.5083	0.6083	0.8083	0.8250	0.9250	0.9750
	c (MLCP2)	0.3028	0.5083	0.6124	0.8010	0.8317	0.9265	0.9618

(MLCP1) stands for the mixed LCP using the transformation of variables (3.5) and (MLCP2) stands for the direct mixed LCP. The blanks correspond to computations not performed by Spence.

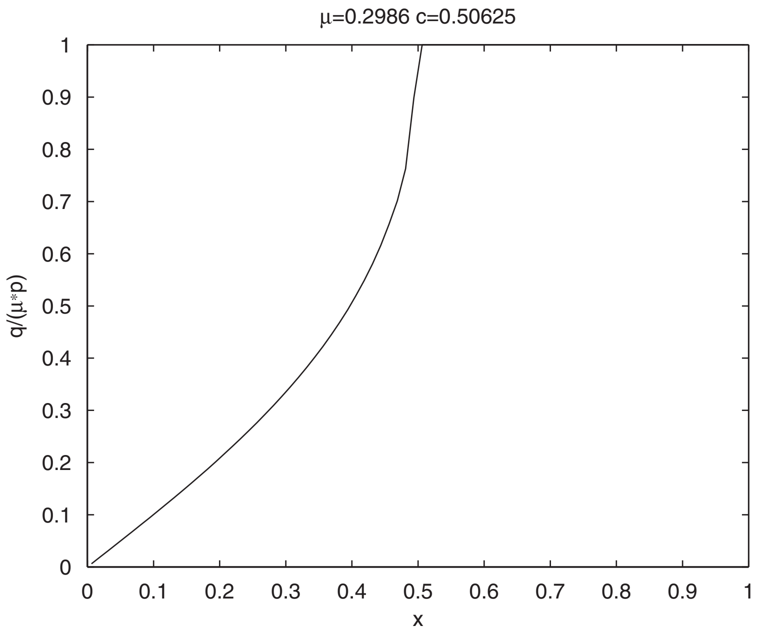

Fig. 2. Ratio $q(x)/\mu p(x)$ in the case $v = 0$.

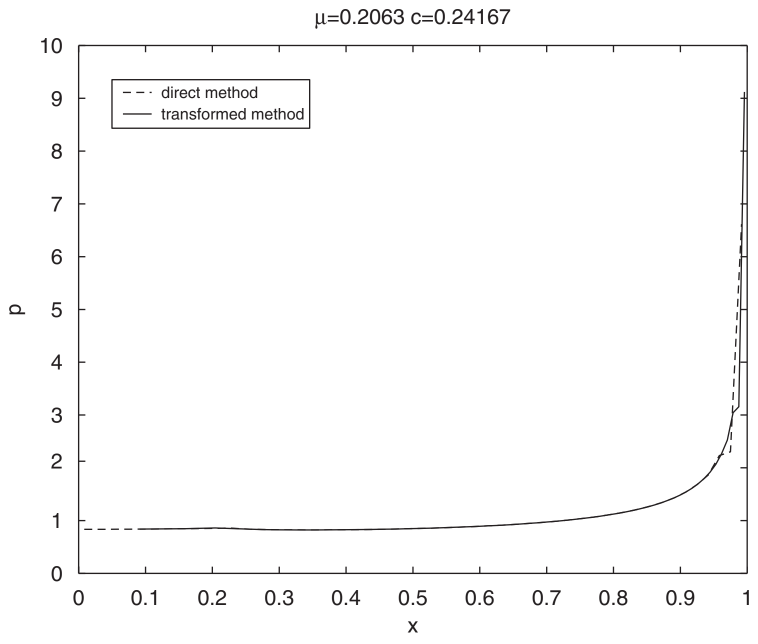

Fig. 3. Normal stress $p(x)$ in the case $v = 0$. The continuous line represents (MLCP1) and dashed line represents (MLCP2).

to the transformation of the dependent variables: it has the effect of mapping a regular grid onto an irregular grid resulting in less precision near $x = 0$ and greater precision near $x = 1$.

The convergence of the method of this paper has been established for a limited range of the parameters v and μ. This is not entirely satisfactory. Now at each iteration of Algorithm 1 it is necessary to solve a linear system. The matrix, $(-\nabla \mathbf{M}(\mathbf{p}^k, \mathbf{q}^k, \mathbf{u}^{*^k}))$, of this linear system has been denoted by $\mathbf{M}_\mathrm{D}$ (or $\mathbf{M}_\mathrm{D}(k)$ to emphasise the dependence on the

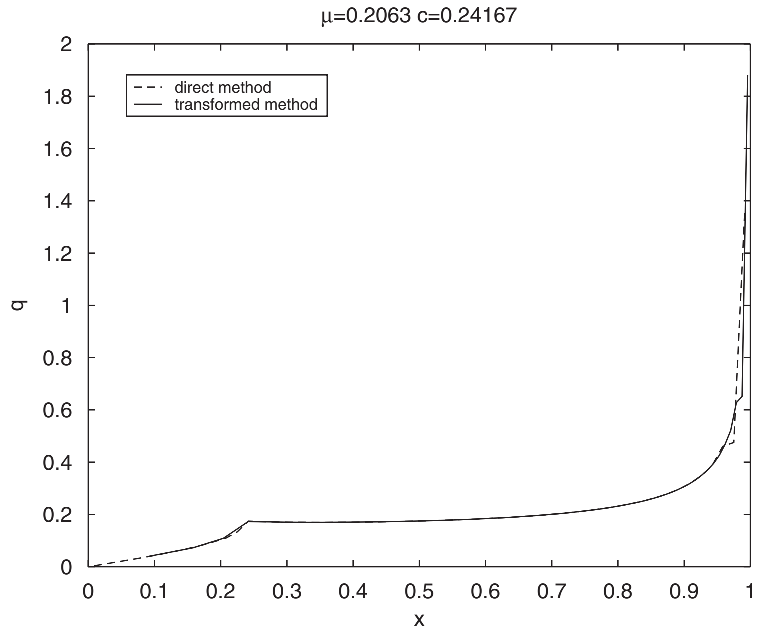

Fig. 4. Shear stress $q(x)$ in the case $v=0$. The continuous line represents (MLCP1) and dashed line represents (MLCP2).

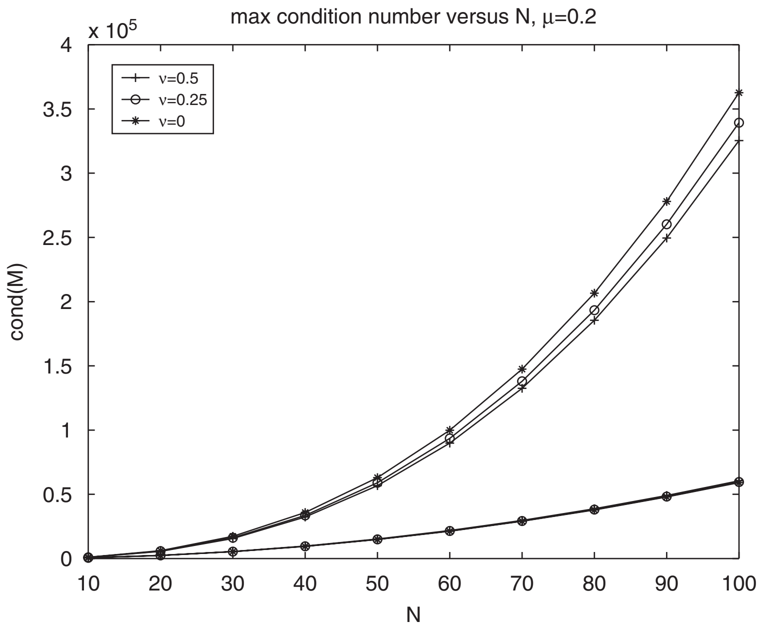

Fig. 5. Condition number of $\mathbf{M}_\mathrm{D}$ and $\tilde{\mathbf{M}}_\mathrm{D}$ for three fixed values of v versus N. The group of curves higher represents $\tilde{\mathbf{M}}_\mathrm{D}$, the other one $\mathbf{M}_\mathrm{D}$.

particular iteration), the corresponding matrix in the untransformed case is denoted by $\tilde{\mathbf{M}}_\mathrm{D}(k)$. Fig. 5 represents the maximum condition number of $\mathbf{M}_\mathrm{D}(k)$, and $\tilde{\mathbf{M}}_\mathrm{D}(k)$, taken over the iterations $1 \leqslant k \leqslant k_f$, where k_f represents the final iteration of the algorithm plotted against N, a measure of the order of the system. The dependence on the parameter v was minimal and no dependence on the parameter μ was observed. However, the dependence of the condition number on the choice of mixed LCP formulation was marked, and it was principally for this reason the discretisation with the transformed variable was chosen.

8. Concluding remarks

This paper treated the Hertz contact problem with friction for an axisymmetric flat punch indenting an elastic half-space. The problem was reformulated, following Noble and Spence [13], as a system of coupled singular Volterra integral equations with a free (unknown) point. After a change of variables a discretisation using product integration was then shown that it could be interpreted as a mixed linear complementarity problem (LCP), the constraint arising from the frictional contact. An algorithm due to Chen and Mangasarian [6] was then implemented and convergence was shown to be dependent upon the invertibility of a certain matrix. This matrix was of block structure and some of these blocks were Toeplitz. This permitted limited convergence analysis to be undertaken. The change of variables prior to discretisation was of course not necessary but did lead to a better condition number for the associated matrix as the numerical results demonstrate. A comparison was also effected with Spence's more indirect approach.

This would appear to be the first time that integral equations have given rise to a complementarity problem. Certainly in the integral equation literature it is most unusual to have coupled Volterra integral equations with a free point. It is therefore in itself not without mathematical interest. This work is, however, not complete. It poses a particular form of an open question which has been known to numerical analysts for at least 30 years. The general open question has been written down recently by Brunner [4] and it is generally associated with the convergence of discretisation methods to their underlying singular integral equation. Interestingly in this paper it needs to be true in order to demonstrate the convergence of the iterates of a Newton-type method for nonlinear equations. Indeed the question of convergence of the solution methods to the system of Abel's equations has not been addressed. The question is difficult both because of the free point and the fact that we are dealing with a system.

Acknowledgement

The third author would like to acknowledge early conversations with Tang Tao and Teresa Diogo and in particular Tang Tao's invitation to the Hong Kong Baptist University, where this work began.

Appendix A.

The following algorithm, proposed by Chen and Mangasarian [6], guarantees a solution to problem (5.3) satisfying

$$\|\hat{\mathbf{M}}(\mathbf{p}, \mathbf{q}, \mathbf{u}^*)\|_\infty \leqslant \varepsilon, \tag{A.1}$$

where $\varepsilon > 0$ is a selected tolerance.

Algorithm 1. (1) *Set tolerance* ε *and define* $\alpha_{\max} = \sqrt{2}/\varepsilon$.

(2) *Initialisation*: *Set* $k = 0$ *and* $\beta = 1/\alpha^0$ *where* $\alpha^0 = \alpha(\mathbf{p}^0, \mathbf{q}^0, \mathbf{u}^{*^0})$ *defined by* (5.6).

(3) *If* $\|\hat{\mathbf{M}}(\mathbf{p}^k, \mathbf{q}^k, \mathbf{u}^{*^k})\|_\infty \leqslant \varepsilon$ *stop.*

(4) *Direction for Newton's method* $\mathbf{d}^k$:

$$\mathbf{d}^k = -(\nabla \mathbf{M}(\mathbf{p}^k, \mathbf{q}^k, \mathbf{u}^{*^k}))^{-1} \mathbf{M}(\mathbf{p}^k, \mathbf{q}^k, \mathbf{u}^{*^k}).$$

(5) *Step size* λ^k:

$$(\mathbf{p}^{k+1}, \mathbf{q}^{k+1}, \mathbf{u}^{*^{k+1}})^{\mathrm{T}} = (\mathbf{p}^k, \mathbf{q}^k, \mathbf{u}^{*^k})^{\mathrm{T}} + \lambda^k \mathbf{d}^k, \quad \lambda^k = \max\{1, e, e^2, \ldots\},$$

such that

$$f(\mathbf{p}^{k+1}, \mathbf{q}^{k+1}, \mathbf{u}^{*k+1}) \leqslant f(\mathbf{p}^k, \mathbf{q}^k, \mathbf{u}^{*^k}),$$

where $e < 1$.

(6) *Parameter update*: *If* $\alpha(\mathbf{p}^{k+1}, \mathbf{q}^{k+1}, \mathbf{u}^{*^{k+1}}) \geqslant \eta\alpha^k$ *set*

$$\alpha^{k+1} = \alpha(\mathbf{p}^{k+1}, \mathbf{q}^{k+1}, \mathbf{u}^{*^{k+1}}),$$

otherwise if $\|\nabla f(\mathbf{p}^{k+1}, \mathbf{q}^{k+1}, \mathbf{u}^{*^{k+1}})\|_2 \leqslant \varepsilon$, *set*

$$\alpha^{k+1} = \eta_1 \alpha(\mathbf{p}^k, \mathbf{q}^k, \mathbf{u}^{*^k}),$$

where $\eta \geqslant \eta_1 > 1$, η *and* η_1 *are given tolerances. If* $\alpha^{k+1} > \alpha_{\max}$, *set* $\alpha^{k+1} = \alpha_{\max}$. *Further let* $\beta^{k+1} = 1/\alpha^{k+1}$, *set* $k := k + 1$ *and go to step* (2).

References

[1] M. Anitescu, J.F. Cremer, F.P. Potra, Properties of complementarity formulations for contact problems with friction, Technical Report 83, Department of Mathematics, The University of Iowa, 1995.

[2] K.E. Atkinson, The numerical solution of an Abel integral equation by a product trapezoidal rule, SIAM J. Numer. Anal. 11 (1974) 97–101.

[3] S.P. Billups, S.P. Dirkse, M.C. Ferris, A comparison of large scale mixed complementarity problem solvers, Comput. Optim. Appl. 7 (1997) 3–25.

[4] H. Brunner, Discretization of Volterra integral equations with weakly singular kernels, in: S. Elaydi, G. Ladas, J. Wu, X. Zou (Eds.), Difference and Differential Equations, Fields Institute Communications, vol. 42, 2004, pp. 419–422.

[5] H. Brunner, The discretization of Volterra functional integral equations with proportional delays, in: S. Elaydi, G. Ladas, J. Wu, X. Zou (Eds.), Difference and Differential Equations, Fields Institute Communications, vol. 42, 2004, pp. 3–27.

[6] C. Chen, O.L. Mangasarian, A class of smoothing functions for nonlinear and mixed complementarity problems, Comput. Optim. Appl. 5 (1996) 97–138.

[7] R.W.G. Cottle, F.L. Giannessi, J.-L. Lions (Eds.), Variational Inequalities and Complementarity Problems, Wiley, New York, 1980.

[8] S.P. Dirkse, M.C. Ferris, The PATH solver: a monotone stabilization scheme for mixed complementarity problems, Optim. Methods Software 5 (1995) 123–156.

[9] N.J. Higham, Accuracy and Stability of Numerical Algorithms, second ed., SIAM, Philadelphia, PA, 2002.

[10] P.A. Holyhead, S. McKee, P.J. Taylor, Multistep methods for solving linear Volterra integral equations of the first kind, SIAM J. Numer. Anal. 12 (1975) 698–711.

[11] L. Johansson, A. Klarbring, The rigid punch problem with friction using variational inequalities and linear complementarity, Mech. Structures Mach. 20 (3) (1992) 293–319.

[12] N. Kikuchi, J.T. Oden, Contact Problems in Elasticity: A study of Variational Inequalities and Finite Element Methods, SIAM Publications, Philadelphia, PA, 1988.

[13] B. Noble, D.A. Spence, Formulation of two dimensional and axisymmetric problems for an elastic half-space, Technical Report 1089, Mathematics Research Center, University of Wisconsin, 1971.

[14] I.N. Sneddon, Fourier Transforms, McGraw-Hill, New York, 1951.

[15] D.A. Spence, The Hertz contact problem with finite friction, J. Elasticity 5 (3–4) (1975) 297–319.

[16] R. Weiss, J.S. Anderssen, A product integration method for a class of singular first kind Volterra equations, Numer. Math. 18 (1972) 442–456.